Using Multivariate Statistical Methods to Detect Fires

Thomas J. McAvoy, James Milke, and *Tekin A. Kunt*
University of Maryland, College Park, Maryland

Abstract

Fire detectors must accurately detect fires, but they should not respond to false alarms. Contemporary smoke detectors sometimes cannot discriminate between smoke and odor sources. These detectors can also be slow in responding to smoldering fire sources. In this paper, a statistical approach for detecting fires based on fusing sensor signals from multiple sensors is presented. The multivariate statistical approach, called principal component analysis, is used to compress the sensor information down to a small number of variables that can be interpreted more easily than the raw sensor signals themselves. Experimental results presented here show that the proposed approach is more accurate than a conventional smoke alarm, particularly for early detection of smoldering fires. However, this new approach does not overcome the problem of false alarms. In spite of this current limitation, the method discussed holds great promise for future fire detection applications.

Introduction

Fire detectors are intended to be sensitive enough to detect fires promptly but not so sensitive that they react to false sources. Contemporary smoke detectors can respond quickly, but generally cannot discriminate between smoke and odor sources. The inability to discriminate between sources is a significant limitation. Data from U.S. fire incidents during the 1980s indicate that 95% of all alarms from smoke detectors were unnecessary.[1]

One solution proposed for minimizing unnecessary alarms without sacrificing prompt activation involves using intelligence along with current detector technology. Some recently developed intelligent detectors move in this direction by incorporating the capacity to correct for background noise, ambient conditions, or changes in detector sensitivity.[2,3] However, these contemporary detectors still are not capable of adjusting even to commonly encountered temporary conditions from tobacco smoke, cooking odors, or aerosol sprays. The next step in the evolution of a smart detector involves incorporating intelligence, possibly with additional sensors, to provide the capability to discriminate between conditions from fire and nonfire sources, without sacrificing response time.[4]

An appreciable amount of effort is being expended by industry to develop odor detection based on an analysis of the response from an array of sensors.[5] For example, applications for smart detectors that have been developed for the food and tobacco industry include process control for products such as coffee and

Fire Technology, Vol. 32, No. 1, 1st Quarter, Jan./Feb., 1996

beer, and quality control evaluations of coffee beans and tobacco blends for cigarettes. Implementing odor detectors for these industrial applications has shown that an accurate assessment of environmental odors is possible as a result of recent developments in sensor technology and analysis techniques.

The feasibility of applying odor detection using metal oxide sensors for fire detection has been demonstrated by Okayama.[6,7] In detecting an odor or a fire, the most desirable situation would involve using completely independent sensors that measure distinct phenomena. Often, however, such independent sensors cannot be found, and one is forced to work with sensors that are not unique and which respond to a number of different phenomena. In this paper, a statistical approach to handling the analysis of measurements from nonunique sensors is discussed.

Teams in the Departments of Fire Protection Engineering and Chemical Engineering at the University of Maryland are conducting ongoing research to determine whether a sufficient distinction in fire signatures can be observed to support developing a smart fire detector. The fire protection engineering team is concentrating on identifying signatures from fire and nonfire sources. The chemical engineering team is applying signal processing techniques, including statistical methods, expert systems, and neural networks, to investigate the sensor response patterns and provide the discrimination capability between fire and nonfire sources. This paper discusses the use of modern, multivariate statistical methods to process signature data from multiple sensors.

In the chemical process industries, distributed control systems routinely collect process data on hundreds, or even thousands, of process variables in real time. The measurements typically involve variables like flow, temperature, pressure, and level. Increasingly, quality measurements such as composition, or product properties, are also being logged. Process operators can be swamped with the sheer volume of information available on a process. In order to cope with this information explosion, multivariate statistical approaches are beginning to be employed effectively to help operators diagnose abnormal operating conditions.[8,9,10] These multivariate approaches, coupled with modern sensing technology, hold promise as a means of detecting fires more accurately. The purpose of this paper is to give an overview of these statistical approaches, and to show how well they work on experimental fire data.

There are two broad types of multivariate statistical methods being used. These methods fall under the headings of principal component analysis (PCA), and partial least squares (PLS).[9] With PCA, one has a set of measurements, $\mathbf{x}$, in which the individual components, x_i, are typically correlated with one another. For example, $\mathbf{x}$ may consist of several temperatures, all of which go up when one goes up. With PLS one has the $\mathbf{x}$ measurements as well as a set of $\mathbf{y}$ variables that are to be predicted from $\mathbf{x}$. For example, the $\mathbf{y}$ variables might be product quality, and the $\mathbf{x}$ variables might be process conditions such as temperature, flow,

pressure, and so on. The result of a PLS analysis is a model to predict **y** from **x**. For the fire detection problem, the PCA methods are more applicable. The remainder of this paper is restricted to these methods.

Multivariate Statistical Methods

PCA dates back to 1901, when it was introduced by Pearson,[11] and to 1933, when the concept of PCA was generalized by Hotelling.[12] PCA uses a set of experimental measurements arranged in a data matrix, **X**. Each row of **X** consists of one set of readings for all m of the x_i's under consideration. The number of rows in **X**, n, equals the number of experimental measurements made. Methods are available to handle the case where some x_i's are missing for particular rows.[9] In chemical process applications, there are usually many more elements in x than there are underlying degrees of freedom in the process. This is due to the correlation of raw measurements. PCA finds linear combinations in the raw measurements that are capable of explaining most of the variation in the raw data.

To understand how PCA works, consider the two-dimensional data set shown in Figure 1. The *o's* represent the collected x_1, x_2 data. Though two variables are measured, it is apparent from the plot that these variables are strongly correlated with one another. When x_1 goes up, so does x_2.

Before discussing the mathematics of PCA, a qualitative description will be given. Suppose that one posed the problem of finding a linear combination of x_1 and x_2 with the maximum possible variance, subject to a magnitude constraint on the linear coefficients. The result of solving this optimization problem would be the first PCA component, the straight line shown in Figure 1. If one used this linear relationship, rather than the raw measurements, then almost all of the varia-

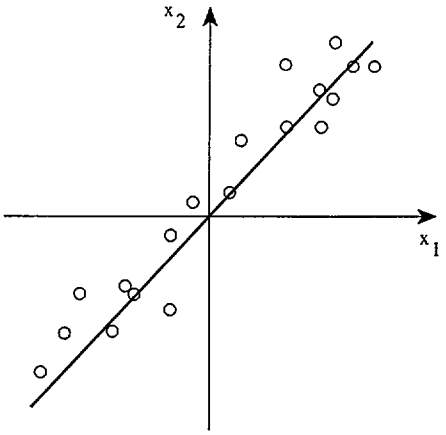


Figure 1. Two dimensional illustration of PCA.

tion in the measurements would be accounted for. In essence, PCA compresses measurements down to a lower dimension, and researchers have found that data interpretation is easier in the reduced space. For two dimensions, the best one can do is reduce the dimensionality to one. In most real applications, PCA achieves a very large reduction in dimensionality, without significant loss of accuracy. For example, on one real data set involving 41 measurements, we were able to attribute 80% of the variation in the data to 4 PCA components.[8]

PCA Mathematics

Mathematically, PCA is based on solving this optimization problem:[13]

$$\max_{p_1} \quad \mathbf{p}' \mathbf{X} \mathbf{p}_1$$

subject to

$$\mathbf{p}'_1 \mathbf{p}_1 = 1 \tag{1}$$

where $\mathbf{X}$ is the matrix of data, $\mathbf{p}_1$ is the set of coefficients, called loadings, to be determined via optimization, and the symbol ' indicates matrix transformation. The solution to Equation 1 gives the direction of maximum variation in the data set under study, that is, the first principal component. After finding the coefficients of the most important direction, $\mathbf{p}_1$, the coefficients of the next most important direction, $\mathbf{p}_2$, for the second principal component are found by subtracting out the contribution of the first direction and solving essentially the same optimization problem. One proceeds to find directions in this manner until a stopping criteria is met, for example, at least 80% of the data variation is explained. Using the singular value decomposition of the data matrix, $\mathbf{X}$,[13] one can solve Equation 1 to determine $\mathbf{p}_1$. Using singular value decomposition, $\mathbf{X}$ can be written as:

$$\mathbf{X} = \mathbf{U} \sum \mathbf{P}' \tag{2}$$

where $\mathbf{U}$ is an $n \times n$ matrix, $\sum$ is an $n \times m$ matrix whose off-diagonal elements are zero, and $\mathbf{P}$ is an m × m matrix. The first column of this $\mathbf{P}$ matrix is $\mathbf{p}_1$, the solution to Equation 1.[13] The remaining columns give the other PCA directions, p_i. The matrix product $\mathbf{U}\sum$ is defined as the score matrix, $\mathbf{T}$, and thus the data matrix can be written as:

$$\mathbf{X} = \mathbf{T}\mathbf{P}' \tag{3}$$

One of the useful properties of the $\mathbf{P}$ and $\mathbf{T}$ matrices is that their column vectors form an orthogonal set. Thus:

$$\mathbf{t}_i\mathbf{t}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$\mathbf{p}_i\mathbf{p}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{4}$$

As the result of the orthogonality properties of **T** and **P, X** can be written as the bilinear product:

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1 + \mathbf{t}_2\mathbf{p}_2 + \ldots\ldots \mathbf{t}_m\mathbf{p}_m \tag{5}$$

and **T** can be calculated from **X** as:

$$\mathbf{T} = \mathbf{XP} \tag{6}$$

Equation 5 is an exact decomposition of **X,** since all m terms are used. PCA approximates **X** with a smaller number of terms, r, where typically $r < m$. For two terms, Equation 5 becomes:

$$\mathbf{X} \cong \mathbf{t}_1\mathbf{p}_1 + \mathbf{t}_2\mathbf{p}_2 \tag{7}$$

As stated earlier, the goal of PCA is to use as few terms as possible but still give an accurate representation of **X.**

To use PCA for detecting faults in chemical plants, one starts by collecting data when the process is running fault-free. This data for normal, good operation is then used to build the PCA model. In this section, it will be assumed that two PCA components give adequate model accuracy for the data. The number of terms to use has to be determined for each application.

One aspect of the two-component PCA model is shown graphically in Figure 2, where the scores are plotted against one another. In Figure 2 the o's represent the good data used to build the model. The ellipse is the confidence limit on the scores for this data, for example, a 95% limit.

Jackson[14] has discussed how this confidence limit can be calculated. The n's in Figure 2 represent new data taken after the model was built. Note that all the measurements contribute to each score point, (t_1,t_2). If a set of scores, designated by x in Figure 2, falls outside the confidence limit, then one can conclude that something has changed in the process. This change may involve a fault, or the fact that the process is operating in a different region than it did when the model was built. Additional information is needed to determine whether a fault has occurred.

Another useful plot, illustrated in Figure 3, involves the squared prediction error (SPE), defined for two PCA components as:

$$SPE = (z \pm t_1 p_1 \pm t_2 p_2)^2 \tag{8}$$

The SPE measures the ability of a reduced dimensional PCA model to predict a set of measurements, $\mathbf{z} = (x_1, x_2, ..x_m)$. First t_1 and t_2 are calculated by applying Equation 6 to z, and then an estimate of z is calculated using Equation 7. Both t_1 and t_2 are linear combinations of all the measurements, $t_i = \mathbf{z}^*\mathbf{p}_i$, with the linear coefficients being the components of the $\mathbf{p}_i$ vectors. Note that for a new data set, $\mathbf{z}$, the t_i's are scalars, while for the original data set, the $\mathbf{t}_i$'s are the vectors of the scores for each row in $\mathbf{X}$—that is, each entry in $\mathbf{t}_i$ gives one score for each row in $\mathbf{X}$.

A confidence limit can also be calculated for the SPE, as discussed by Jackson and Mudholkar.[15] Their equations for the confidence limit are given in the Appendix.

The *o*'s, *n*'s, and *x*'s have the same meaning in this plot as in Figure 2. To use this plot, one would calculate the SPE for new data being measured. As long as the SPE remains below its confidence limit, then one can conclude that the new data agrees with the PCA model. Since the PCA model is based on good operation, the implication of an SPE below its confidence limit is that the new data indicates fault-free operation. If the confidence limit is exceeded, as indicated by the point labeled *x*, then the model is no longer accurate, and a fault is suspected, particularly if the operating region has not changed.

The score and SPE plots have proven to be very useful in detecting faults in chemical processes. Both plots represent a very large compression of informa-

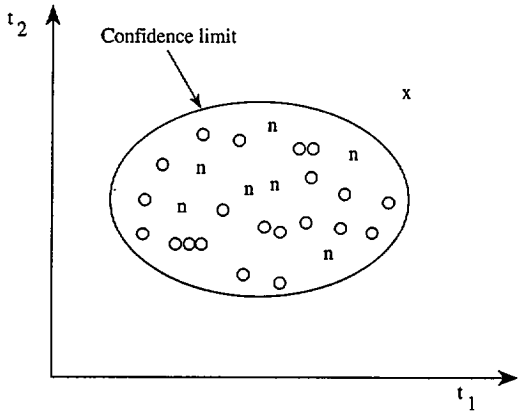


Figure 2. Scoreplot.

tion. Process operators do not have to look at individual process measurements to detect a problem. Rather, they can look at a few diagrams, such as Figures 2 and 3, and through them monitor all process measurements. It should be emphasized that each plot reflects all of the measurements being made.

Modern Sensor Technology

PCA analysis is most effective when one has measurements from many sensors at once. A much greater compression of information occurs when 50 sensors can be reduced to a few scores than it can when only two sensors are involved. Also, use of many sensors may lead to better separation of classes, and therefore, more accurate classification—that is, fire detection. Today, scientists are producing arrays of microsensors on computer chips that can measure many hundreds of readings simultaneously. In this paper, it is impossible to give a thorough overview of the status of sensor research today. Rather, a brief description of one sensor system that we are testing for its fire detection ability is described.

Figure 4 shows a blowup of one sensor element[16] of such a system. This device consists of a micromachined bridge, which supports these three successive layers separated by insulation: first, a polysilicon heater; second, an aluminum hot plate with electrical connections for four-point temperature measurement; and third, open aluminum contact pads to make electrical contact to a deposited overlayer. The sensor is based on measuring conductance changes of the deposited semiconducting oxides at elevated temperatures. This device is fabricated using a combination of CMOS device fabrication and post-CMOS processing steps. The use of CMOS chips offers the advantages of high reliability, low cost, and simple integration of on-chip electronics for multiplexing

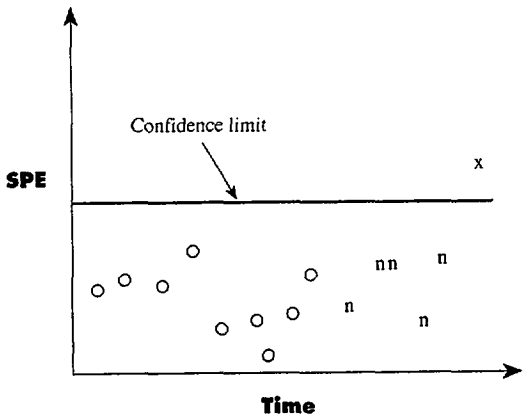


Figure 3. Squared prediction error plot.

and signal processing.

As shown in Figure 4, the dimension of the element is ~200 microns square. Thus, assuming that the associated signal processing and power circuitry each occupy the same space as the sensor, in one square inch over several hundred individual sensors can be placed in an array. By coating each sensor with a different semiconducting material, several hundred different readings can be made. In an actual sensor, it would no doubt be uneconomical to use this many different coatings, but this number does indicate the diversity of possibilities. In addition, the fabricators of the chip shown in Figure 4 have developed an approach to cycling the temperature of each sensor to further enhance its selectivity.[16]

The net results of both the number of sensors and the possibilities to manipulate the temperature profile lead to an information-rich and complex set of sensor readings. It is precisely for such a complex signal that PCA holds its greatest promise. In our fire experiments, we made a total of six measurements. Many more variables can be measured in the future using sensor arrays, and in this way, the potential for enhanced fire detection will be magnified greatly.

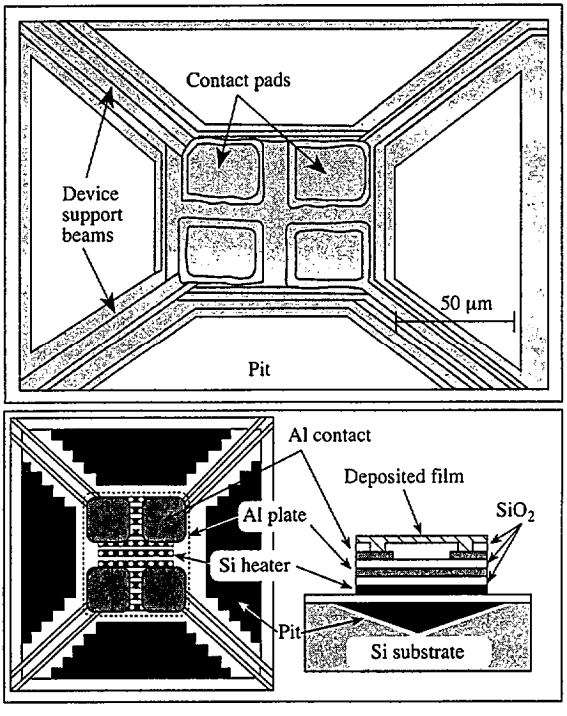
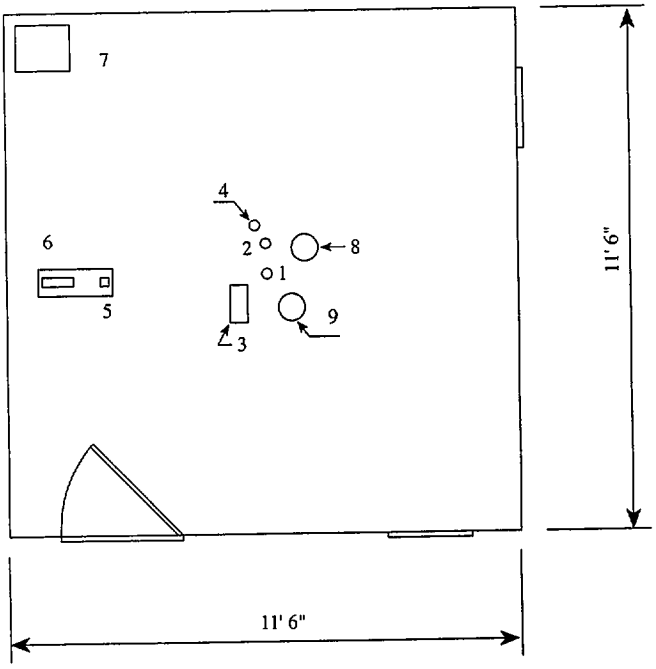


Figure 4. Thin film gas sensor taken from Cavicchi *et al.*[16]

Large-Scale Experimental Program

Large-scale experiments have been conducted to determine whether the trends identified in small-scale experiments[17] also existed in large-scale environments. Signatures from a wide variety of fire and environmental sources were monitored, and sensor response patterns were explored.

The large-scale experiments were conducted in a 3.6 m × 3.6 m room with a height of 2.4 m.[18] Measurements included temperature; mass loss of the fire sources; CO, CO_2, and O_2 concentrations; light obscuration; and the voltage output from two metal oxide sensors (Taguchi models 822 and 880). In addition, two commercial smoke detectors (one photoelectric and one ionization) were mounted on the ceiling, at the center of the room. For a diagram of the room, including the relative locations of the sensors, see Figure 5.



1. Thermocouple tree
2. Taguchi sensor (TGS 822)
3. Taguchi sensor (TGS 880)
4. 3/16" copper sampling tube
5. Photocell
6. Helium-neon laser
7. Load cell
8. Ionization smoke detector
9. Photoelectric smoke detector

Figure 5. Diagram of test room.

The metal oxide sensors responded to the presence of oxidizable gases and environmental odors respectively. Mass loss measurements were used to estimate the yield fractions of the signatures from the fire sources. Because the tests were conducted in an unconditioned space, data were collected for at least two minutes before introducing any source, in order to document ambient conditions. This ambient data is used to build a PCA model under the assumption that it represents normal, nonfire operation. Table 1 gives a summary of the variety of sources used to generate conditions within the room. Again, these were intended to be representative of residential fire and nuisance sources.

Flaming liquid tests were conducted by placing 50 ml of the sample in a precooled metal container ignited by a match. The container was cooled before the tests to limit evaporation of the liquid before initiating flaming. Tests with flaming solids involved placing the fuel in an aluminum pan, then igniting the fuel with a match. Tests with pyrolyzing solids were conducted by placing the fuel in an aluminum pan on a preheated hotplate.

The tests involving the environmental sources were conducted using several approaches, depending on how the product is typically used in a residence. One approach consisted of dispersing the product throughout the room, including water mist, cigarette smoke, and household aerosol products. Alternatively, solid and liquid products, such as bleach, nonacetone nail polish remover, boiling liquids, coffee, and toast were located at floor level in the center of the room. The test with toast was conducted by placing the bread in a toaster that was kept "on" throughout the test. Tests with coffee used both fresh coffee grounds and brewed coffee.

Applying PCA to Fire Experiments

The fire experiments consist of 87 tests broken down as follows: 34 flaming, 16 smoldering, and 37 nuisance cases. Six of the sensor readings recorded during the tests were used to develop the PCA model. These were CO, CO_2, two Taguchi sensors (T880, T882), temperature, and light obscuration. Before beginning any test, data for all six sensors were recorded. The data for each sensor were scaled to zero mean and unit variance. This data was taken to be that for normal, good operation, and a PCA model was built using it. The results are given in Table 2.

Three PCA components explained almost 76% of the variability of the normal data, so three components were used in subsequent analysis. Subsequent test data were scaled using the means and variances of the normal data.

The SPE was used to flag abnormal situations, with its confidence limit being set at 99.5%. When three consecutive SPEs exceeded the 99.5% confidence limit, an abnormal situation was declared. To distinguish among flaming fires, smoldering fires, and nuisances, it was necessary to examine the scores. Table 3 gives the complete scores resulting from the PCA analysis.

The scores shown are those that occurred when an abnormal situation was

TABLE 1
Test Sources

Heated Fuels			Environmental
Liquid	**Solid**	**Gas**	**Sources**
heptane, 1-proponal, methanol, toluene, vegetable oil[1]	paper, cotton, polystyrene, pine, cardboard, cheesecloth, toast[2]	propane	propane, aerosols (disinfectant, furniture polish, cooking spray, hair spray), nail polish remover, ammonia-based window cleaner, bleach, water mist, boiling water, toast, cigarette smoke, coffee

1. boiling only
2. smoldering only

declared. Also shown is a comparison between the detection times for the PCA model and an ordinary smoke detector. The PCA approach missed detecting two smoldering cases. The smoke detector failed to detect 8 smoldering and 8 flaming fires. After studying the scores in Table 3, these three rules were developed to classify an abnormal event: If $t_3 > 5$, then there is a flaming fire. If $-8 < t_2 < 0$, then there is a smoldering fire. Else, there is a nuisance.

It should be noted that these rules are specific to the fire tests that were conducted. If data are generated on other systems, then it should be possible to develop rules in a manner similar to those used here.

The rules developed have some physical interpretation behind them. An examination of the loadings in Table 2 shows that CO_2 contributes the most to t_3. Large values of CO_2 result in flaming fires, and these large values lead to large values of t_3. This is why the first rule involving t_3 was developed. The second "if" statement was developed from studying the score plots, and it was more difficult to write down, since the nuisance and smoldering cases tended to overlap. Figure 6 shows a projection of the scores in Table 3 onto the t_1–t_2 plane for the smoldering and nuisance cases. Note that the various cases are intermingled, and it is not possible to separate them completely with simple rules. Our rules successfully flagged 14 out of 16 smoldering cases, but 10 of the 37 nuisance cases were misclassified as smoldering fires. Thus, there were 2 errors and 10 false alarms.

In addition to the benefit of detecting smoldering fires with a greater frequen-

TABLE 2
Principal Component Model

Percent Variance Captured by PCA Model			
Principal Component	Eigenvalue	% Variance	Cumulative % Variance
1	1.8700	31.1660	31.1660
2	1.4766	24.6096	55.7756
3	1.2086	20.1427	75.9183
4	0.6578	10.9630	86.8813
5	0.4678	7.7963	94.6776
6	0.3193	5.3224	100.0000

PCA Loadings			
	p_1	p_2	p_3
T822	0.5119	–0.3136	0.3216
CO	–0.4875	–0.2472	–0.1623
CO2	–0.0048	–0.1294	0.8433
Light Obsc.	–0.5605	0.1317	0.3273
T880	0.1463	–0.7189	–0.2204
Temperature	–0.4059	–0.5381	0.0590

cy than the commercial detectors, the PCA approach resulted in reduced detection times in all flaming fires and in all the smoldering fires. Table 4 presents a comparison of detection times. As indicated in the table, the mean detection time for flaming fires is reduced 45 s, representing an average decrease of 57%. In one case, the detection time for one flaming fire was reduced from 260 s to 16 s. The average reduction in detection time for smoldering fires was 245 s, with an average decrease of 30%.

Clearly, 2 errors and 10 false alarms out of 87 cases is not acceptable. To improve this ratio, more sensors need to be used. They should be chosen so that the scores for the nuisance and flaming cases separate and an accurate distinction can be made. As discussed, sensor research has progressed to the point where several hundred different gas sensors can be put on a chip. While exactly what sensors to use remains an open question, such a chip, coupled with multivariate statistical approaches, holds great promise as a means of achieving improved fire detection.

TABLE 3
Detection Times for Conventional Smoke Alarm

		Detection Time (min)		Scores		
File	Type[a]	Detector	PCA Model	t_1	t_2	t_3
72904	s	11:31	11:58	1.82	–1.37	1.44
72906	s	14:06	9:36	3.27	–2.35	1.84
72702	s	13:40	11:54	2.21	–2.67	1.11
72505	s	13:46	8:14	0.35	–4.74	0.05
72507	s	12:38	9:20	2.70	–7.32	1.72
80102	s	13:58	10:48	3.00	–5.70	1.49
80806	s	15:24	12:22	3.98	–5.63	1.70
80905	s	14:24	11:18	2.26	–4.04	1.58
72902	s	—	12:40	4.13	–3.62	2.39
72602	s	—	10:52	1.43	–2.31	0.62
80402	s	—	11:18	1.73	–5.64	–0.51
72502	s	—	10:58	0.77	–4.35	0.38
80302	s	—	11:14	2.49	–5.02	0.85
80502	s	—	10:52	1.92	–3.92	2.06
72709	s	—	—	—	—	—
72506	s	—	—	—	—	—
72710	f	—	0:28	0.84	–0.67	6.25
72910	f	—	0:14	0.12	–2.08	6.22
80804	f	—	0:12	–0.84	–1.25	9.32
80505	f	—	0:08	1.44	–3.78	17.01
80507	f	—	0:10	1.26	–2.20	13.66
80508	f	—	0:06	0.57	–3.60	21.89
80906	f	—	0:10	0.54	–5.89	29.62
80306	f	—	0:08	–0.62	–2.67	17.06
72501	f	0:16	0:08	–0.08	–3.43	23.67
72601	f	0:14	0:08	0.12	–0.69	33.13
72607	f	1:30	0:18	1.96	–4.74	13.14
72701	f	0:16	0:08	–1.25	–5.24	15.98
72704	f	0:58	0:24	0.22	–1.18	8.07
72705	f	0:34	0:20	1.50	–4.63	7.27
72711	f	0:58	0:12	0.21	–0.04	5.98
72901	f	0:22	0:08	–0.02	–3.53	34.18
72905	f	0:40	0:08	–1.14	0.46	9.57
72907	f	1:38	0:10	0.02	1.28	11.62

[a] s: smoldering, n: nuisance, f: flaming, —: no detection

TABLE 3, CONT.

		Detection Time (min)		Scores		
File	**Type**[a]	**Detector**	**PCA Model**	t_1	t_2	t_3
80101	f	0:22	0:08	0.27	–4.26	6.06
80104	f	0:24	0:10	1.19	–2.34	8.47
80105	f	1:50	0:10	–1.31	–2.54	11.80
80108	f	0:20	0:08	0.84	–3.85	16.00
80301	f	0:20	0:08	0.78	–3.66	24.90
80304	f	0:20	0:10	–1.08	–4.63	13.98
80401	f	0:14	0:08	–2.33	–5.45	31.65
80501	f	0:20	0:08	–2.13	–3.48	23.13
80801	f	0:22	0:08	0.71	–0.05	32.18
80802	f	3:40	0:38	–0.15	–0.49	6.51
80803	f	0:32	0:12	–0.54	–3.32	26.71
80805	f	1:20	0:08	–0.65	–3.06	21.44
80808	f	0:40	1:10	–1.38	–4.35	26.26
80809	f	4:20	0:16	–0.27	–2.85	8.23
80810	f	2:08	0:12	–1.42	–2.26	10.47
80901	f	0:22	0:08	0.08	–4.54	25.83
72911	n	0:10	1:04	2.46	–1.61	1.05
72608	n	0:10	0:10	35.27	–79.21	–9.41
80410	n	0:16	0:10	33.41	–81.11	–7.74
80407	n	16:38	9:24	0.55	–5.09	1.61
72708	n	21:02	20:20	1.30	–3.07	–0.39
72503	n	—	0:08	28.54	–56.82	–0.98
72603	n	—	0:08	27.66	–55.60	–0.64
72604	n	—	0:10	34.41	–80.81	–6.58
72605	n	—	0:48	3.14	–6.44	2.47
72609	n	—	1:48	3.24	0.04	1.04
72703	n	—	0:30	23.94	–32.14	2.92
72706	n	—	5:46	2.12	–3.05	0.80
72903	n	—	0:08	29.10	–52.29	–1.26
80303	n	—	0:08	29.46	–57.56	–1.82
80305	n	—	8:30	6.76	–9.24	1.52
80307	n	—	8:30	2.04	–10.74	–1.36
80403	n	—	0:10	29.37	–61.04	–1.64
80404	n	—	3:40	1.65	–2.92	1.11

[a] s: smoldering, n: nuisance, f: flaming, —: no detection

TABLE 3, CONT.

		Detection Time (min)		Scores		
File	**Type[a]**	**Detector**	**PCA Model**	t_1	t_2	t_3
80405	n	—	1:08	0.93	–5.24	1.72
80406	n	—	1:18	3.04	–6.13	0.92
80408	n	—	7:56	1.51	–5.82	0.80
80506	n	—	7:24	7.86	–10.07	1.75
80807	n	—	4:34	2.72	–4.81	0.13
80103	n	—	0:08	28.88	–54.67	–0.39
80106	n	—	0:08	27.38	–55.37	–1.21
80409	n	—	0:10	24.56	–55.39	–2.53
80503	n	—	0:10	31.03	–60.35	–1.85
72909	n	—	—	—	—	—
80107	n	—	—	—	—	—
80504	n	—	—	—	—	—
80902	n	—	—	—	—	—
80903	n	—	—	—	—	—
80904	n	—	—	—	—	—
72504	n	—	—	—	—	—
72706	n	—	—	—	—	—
72707	n	—	—	—	—	—
72908	n	—	—	—	—	—

[a] s: smoldering, n: nuisance, f: flaming, —: no detection

Summary

As a result of the experimental effort, an early fire detector consisting of an array of gas sensors appears feasible, with discrimination provided by a principal component analysis (PCA) of the sensor responses. It has been shown that discrimination among a flaming fire, smoldering fire, and a nuisance source can be done effectively using the PCA factors. However, many questions must be addressed before applying this technology as a means of early fire detection.

Due to the small number of sensors used, namely six, a number of false alarms occurred for nuisance sources, which were classified as smoldering fires. Adding more sensors should help to alleviate this problem. Additional research is required to optimize the number and types of sensors to be included in the array, while still providing the desired level of sensitivity and discrimination ability. In addition, the data has been acquired from experiments conducted with one type of source, for example, a flaming source without a nuisance source being present.

TABLE 4
Time to Detection

	Flaming Fires	Smoldering Fires
Total	34	16
Number of fires undetected—commercial	8	8
Number of fires undetected—PCA	0	2
Average reduction in detection time (s)	45 (57%)	245 (30%)
Range of reduction in detection time (s)	6–244 (41–94%)	182–332 (20–40%)

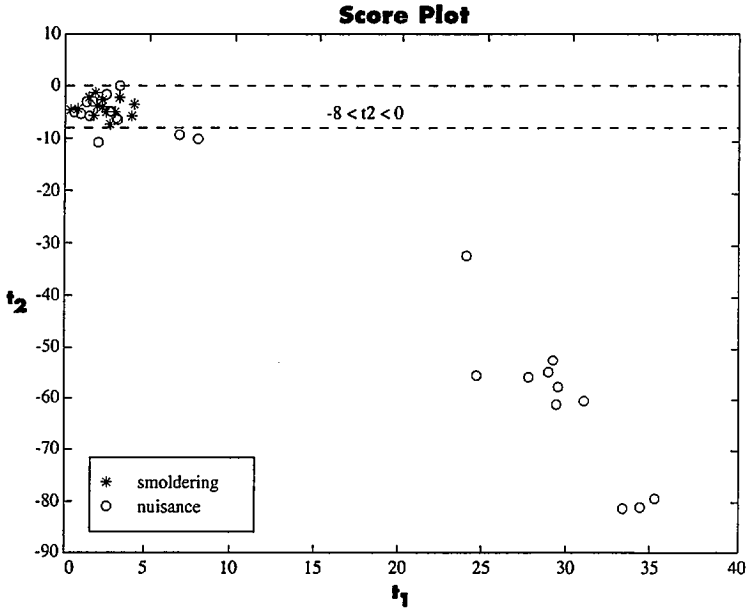


Figure 6. Projection of scores for nuisance and flaming experiments, t_1–t_2 plane.

Additional experiments are needed to assess the potential for a nuisance source to mask a flaming or smoldering source.

Acknowledgment

This project is supported by the Building and Fire Research Laboratory of the National Institute of Standards and Technology (NIST). Dr. William Grosshandler is the technical monitor.

Nomenclature

c	Defined by Equation A-5
h_0	$1 - (2\theta_1\theta_3/(3\theta_2^2))$
n	Number of PCA components used in model
m	Total number of measurements
$\mathbf{p}_i$	Column vector of loading coefficients for PCA factor i
$\mathbf{P}$	Matrix of loading vectors. $P = [p_1 p_2,...,p_n]$
Q_α	Confidence limit on SPE at level α given by Equation A-6.
t_i	Score value for a new measurement, z. $t_i = z*p_i$
$\mathbf{t}_i$	Column vector of scores equal to $X*p_i$
$\mathbf{T}$	Matrix of score vectors. $T = [t_1 t_2,...t_n]$
$\mathbf{U}$	Matrix calculated in singular value decomposition of X
x_i	Individual measurement of variable i
$\mathbf{x}_i$	Column vector of measurement x_i from different experiments
$\mathbf{X}$	Matrix of data used to build PCA model. $X=[x_1 x_2,...x_m]$
$\mathbf{z}$	Row vector of new measurements. $z = (x_1 x_2,...x_m)$
α	Confidence level
θ_i	Parameter defined by Equations A-1 to A-3
$\sigma_{i,i}$	Diagonal element of Σ

References

1. Hall, J. R., "The Latest Statistics on U.S. Home Smoke Detectors," *Fire Journal,* Vol. 83, No. 1 (1989), pp. 39–41.

2. Bukowski, R. W. and O'Laughlin, R. O., *Fire Alarm Signaling Systems.* Quincy, Mass: NFPA, 1994.

3. Thuillard, M., "New Methods for Reducing the Number of False Alarms in Fire Detection Systems," *Fire Technology,* Vol. 30, No. 2 (1994), pp. 250–268.

4. Grosshandler, W. L., *An Assessment of Technologies for Advanced Fire Detection.* Gaithersburg: NIST, 1992.

5. Gardner, J. W., Shurmer, H. V. and Tan, T. T., "Application of an Electronic Nose to the Discrimination of Coffees," Sensors and Actuators B, Vol. 6 (1992), pp. 71–75.

6. Okayama, Y., "Approach to Detection of Fires in Their Very Early Stage by Odor Sensors and Neural Net," *Proceedings of the 3rd International Symposium of Fire Safety Science* (1991), p. 955–964.

7. Okayama, Y., Ito, T., and Sasaki, T., "Design of Neural Net to Detect Early Stage of Fires and Evaluation by Using Real Sensors' Data," *Proceedings of the 4th International Symposium on Fire Safety Science,* (1994), Ottawa.

8. Dong, D., and McAvoy, T., "Nonlinear Principal Component Analysis Based on Principal Curves and Neural Networks," *Computers and Chemical Engineering,* Vol. 20 (1996), pp. 65–78.

9. Kresta, J., Macgregor, J., and Marlin, T., "Multivariate Statistical Monitoring of Process Operating Performance," *Canadian Journal of Chemical Engineering,* Vol. 69 (1991), pp. 35–47.

10. Nomikos, P., and Macgregor, J., "Monitoring Batch Processes Using Multi-way PCA," *AIChE Journal,* Vol. 40 (1994), pp. 1361–1375.

11. Pearson, K., "On Lines and Planes of Closest Fit to Systems of Points in Space," Phil. Mag., Vol. 2 (1901), pp. 559–572.

12. Hotelling, H., "Analysis of a Complex of Statistical Variables into Principal Components", Journal of Educational Psychology, Vol. 24 (1933) pp. 417–441, and 498–520.

13. Jolliffe, I., *Principal Component Analysis.* New York: Springer-Verlag, 1986.

14. Jackson, J., "Principal Components and Factor Analysis: Part I-Principal Components," *Journal of Quality Technology,* Vol. 12 (1980), pp. 201–213.

15. Jackson, J., and Mudholkar, G., "Control Procedures for Residuals Associated with Principal Component Analysis," *Technometrics,* Vol. 21 (1979), pp. 341–349.

16. Cavicchi, R. Suehle, J., Chaparala, P., Poirier, G., Kreider, K., Gaitan, M., and Semancik, S., "Micro-hotplate Temperature Cointrol for Sensor Fabrication, Study, and Operation," *Proceedings of the Fifth International Meeting on Chemical Sensors,* (July 1994), Rome, Italy.

17. Denny, Samuel, "Development of a Discriminating Fire Detector for Use in Residential Occupancies," Report FP 93-07, M.S. Thesis, College Park, Fire Protection Engineering, University of Maryland, December 1993.

18. Hagen, B. C., "Evaluation of Gaseous Signatures in Large-Scale Test," Report FP 94-05, M.S. Thesis, College Park, Fire Protection Engineering, University of Maryland, December 1994.

Appendix

This Appendix gives the equations from Jackson and Mudholkar[15] that can be used to calculated confidence limits on the SPE. First, define

$$\theta_1 = \sum \sigma_{i,i} \tag{A-1}$$

$$\theta_2 = \sum \sigma_{i,i}^2 \tag{A-2}$$

$$\theta_3 = \sum \sigma_{i,i}^3 \tag{A-3}$$

where $\sigma_{i,i}$ is a diagonal entry of Σ, and the summations go from the number of principal components used plus 1, $n + 1$, to the total number of measurements, m. Let $h_0 = 1 \pm (2\theta_1\theta_3 / (3\theta_2^2))$, then the quantity $(Q / \theta_1)^{h_0}$ is approximately normally distributed as:

$$(Q / \theta_1)^{h_0} \approx N\left|1 + \theta_2 h_0 (h_0 \pm 1) / \theta_1^2, 2\theta_2 h_0^2 / \theta_1^2\right| \tag{A-4}$$

$$c = \frac{\theta_1\left[(Q / \theta_1)^{h_0} \pm 1 \pm \theta_2 h_0 (h_0 \pm 1) / \theta_1^2\right]}{\sqrt{2\theta_2 h_0^2}} \tag{A-5}$$

Then for a fixed type I error α, the confidence limit on the SPE can be approximated as:

$$Q_\alpha \cong \left[c_\alpha \sqrt{2\theta_2 h_0^2} / \theta_1 + 1 + \theta_2 h_0 (h_0 \pm 1) / \theta_1^2\right]^{1/h_0} \tag{A-6}$$